基于抗体网络的邮件过滤器设计

匡 胤[1]，黄迪明[2]

(1. 内江师范学院计算机科学系 四川 内江 641002; 2. 电子科技大学计算机科学与工程学院 成都 610054)

【摘要】抗体网络作为一种新型的基于免疫原理的神经网络模型，已有实验验证了其具有良好的无监督竞争学习能力。但对抗体网络的研究目前还集中在原理介绍和实验验证上，没有将其应用在实际工程问题中的先例。在保留抗体网络的结构自动生成，基于克隆选择、变异机制的无监督竞争学习等优点的同时，对抗体网络的初始化、抗体的表示方式、网络结构的更新等方面作了适当的改进。在此基础上设计的邮件过滤器，和传统的邮件过滤器相比，实验结果表明其具有自适应能力好、准确性高等优点。

关 键 词 垃圾邮件; 邮件过滤器; 抗体网络; 克隆选择; 人工神经网络
中图分类号 TP18 文献标识码 A

A Design of Mail Filter Based on Antibody Network

KUANG Yin[1]，HUANG Di-ming[2]

(1. Department of Computer，Neijiang Teachers College Neijiang Sichuan 641002;

2．School of Computer Science & Engineering，Univ. of Electron. Sci.& Tech. of China Chengdu 610054)

Abstract The Antibody Network (ABNET), which is a new Artificial Neural Networks (ANN) based on immune principle, has been proved to have good ability of unsupervised and competitive learning in experiments. But there is no precedent to use ABNET in practical engineerings because the present researches about it are still focusing on principle and experiments. While the good qualities of ABNET are reserved, such as structure generated automatically, unsupervised and competitive learning based on clone selection and mutation,the way of initiating ABNET, expressing a antibody, and updating structure is improved properly. The mail filter based on improved ABNET is better at adaptation and accuracy in experiments, compared with traditional mail filters.

Key words spam; mail filter; ABNET; clonal selection; ANN

人工免疫系统(Artificial Immune System，AIS)和人工神经网络(Artificial Neural Networks，ANN)是目前计算智能的两个研究热点。自从Hoffmann率先将人工免疫理论应用到人工神经网络的设计中以后，人工免疫系统极大地促进了神经网络的研究，陆续有学者提出新的基于免疫的神经网络模型，这些模型在不同的工程领域得到广泛应用[1-5]。文献[6]提出了一种基于人工免疫原理的竞争神经网络模型——抗体网络(Antibody Network，ABNET)。为了评估其性能，抗体网络被应用到三个不同的问题中进行验证。实验结果表明，抗体网络在某些特定的能力上超过实验中的其他对比模型[6]。目前见诸文献的对抗体网络的研究集中在原理介绍和实验验证上[6-8]，还没有应用在工程实际问题中的先例。

1 垃圾邮件(Spam)与邮件过滤技术

垃圾邮件一直是Internet的顽症之一。垃圾邮件不仅浪费网络资源，同时还可能造成巨大的社会负面影响。因此，反垃圾邮件技术的研究势在必行。

人工免疫系统在工程领域的应用研究，为防范垃圾邮件带来了新的希望。目前，基于人工免疫系统的邮件过滤技术逐渐得到学术界的重视，国外已有基于人工免疫原理的反垃圾邮件产品问世。

2 过滤器的设计

2.1 过滤器的系统模型

过滤器的工作原理如图1所示[9]。基因库模拟人体的骨髓，以恒定的速率随机产生初始B细胞，并加入

收稿日期：2004 - 05 - 23
基金项目：四川省科技厅重点攻关项目(03GG006-021)；四川省教育厅青年基金资助项目(2005B043)
作者简介：匡 胤(1974 -)，男，硕士，讲师，主要从事网络安全和智能计算方面的研究.

抗体网络中。B细胞的生命值同时也以一定速率减少，当其为0时，B细胞将会死亡，自动从网络中被剪枝。这种机制体现了抗体网络竞争学习的特点，并且可以保证B细胞种群的多样性和稳定性。

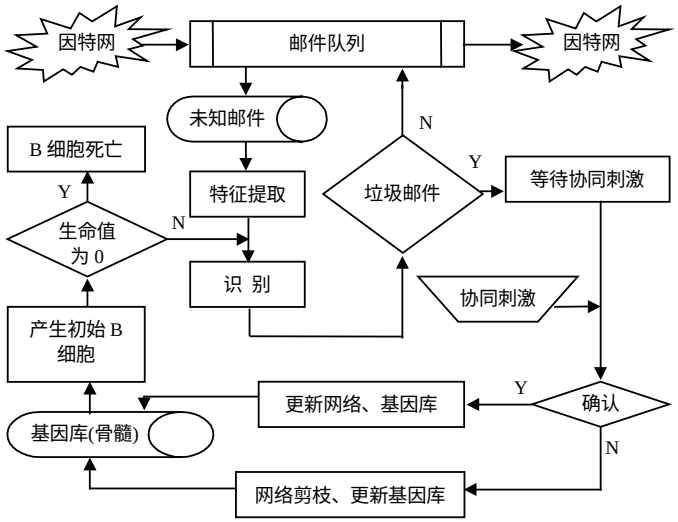

图1 邮件过滤器系统流图

2.2 基因库的构造

根据收集到的垃圾邮件样本，分别对主题(Subject)字段、发送人(Sender)字段和正文(Message)字段进行特征提取，去掉重复项和一些常用关键词(如and，of等)后，形成三个由特征项构成的集合，这三个集合共同构成基因库。

2.3 抗原及B细胞的构造

抗原及B细胞的构造基于向量空间模型，两者都定义为：

vector =< subject，sender，message >

subject=< term 1，term 2，…，term *i* >

sender =< term 1，term 2，…，term *j* >

message =< term 1，term 2，…，term *k* >

特征向量的长度可变且与特征项的顺序无关。两者不同之处在于：抗原的特征向量从未知邮件中提取特征项构成；B细胞的特征向量是从基因库的三个集合中分别随机挑选特征项构成。

2.4 抗体网络的结构

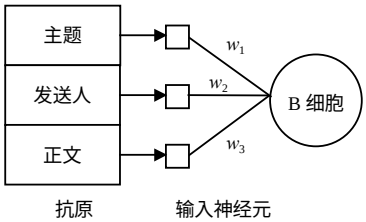

图2 抗体网络的结构

网络的拓扑结构用权值矩阵来表示和存储。每个输入神经元对应矩阵的一行，每个B细胞对应矩阵的一列。网络中有三个输入神经元，分别顺序输入抗原的主题、发送人、正文三个向量，这三个向量与B细胞的对应的三个向量分别计算相似性(相同特征项个数/较短向量的长度)，得到w_1、w_2、w_3三个值w_1、w_2、$w_3 \in [0, 1]$，这三个值作为输入神经元与B细胞的连接权值保存在矩阵中。B细胞和抗原的亲和力定义为affinity:=$(w_1+w_2+w_3)/3$。图2表示只有一个B细胞的抗体网络，图3表示对应的权值矩阵。

2.5 两个重要的算法模块

2.5.1 未知邮件的识别

抗体网络初始化完成后，即可以对未知邮件进行识别，识别过程描述如下：

```
PROCEDURE   is_spam(mail_vector)
BEGIN
  max_ affinity:=0.0；j:=0；/* j指示亲和力最大的B细胞*/
  FOR i:=1 TO B细胞当前种群数  DO
```

BEGIN
计算第i个B细胞与mail_vector的权值w_1、w_2、w_3；
w_1、w_2、w_3赋给权值矩阵的第i列；
IF ((affinity:= (w_1+w_2+w_3)/3)≥ε　AND　(affinity> max_ affinity))　/*ε为预定阈值*/
THEN　BEGIN　j:=i；max_ affinity:= affinity；END；
END；/*未知邮件识别完成*/
IF (j=0) THEN RETURN “normal”；
ELSE RETURN “spam”；
END.

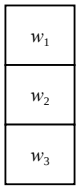

图3　权值矩阵

2.5.2　抗体网络和基因库的更新

如果过滤器将某邮件识别为垃圾邮件，将会等待用户反馈回来的协同刺激，以便对抗体网络和基因库更新。算法描述如下：

PROCEDURE　update(message) /* message为协同刺激信号 */
BEGIN
IF (message= “spam”)
THEN BEGIN
该邮件特征项加入基因库；
增加第j个B细胞的生命值； /* 鼓励识别效果最好的B细胞 */
克隆第j个B细胞，特征项发生变异的比例为max_ affinity；
为新产生的B细胞在权值矩阵中增加一个列$(0.0，0.0，0.0)^T$；
END；
ELSE BEGIN
从基因库中删除构成该邮件的特征项；
将所有识别该邮件为垃圾邮件的B细胞剪枝，并从矩阵中删除对应列；
END；
END.

3　实验及结果分析

在设计实验时，重点验证过滤器的学习和自适应能力。垃圾邮件样本来自一个专业反垃圾邮件组织SpamAssassin。该组织2005年3月提供的邮件样本库中包含6 047个样本，其中垃圾邮件1 897例，正常邮件4 150例。正常邮件样本中包含2 500例非常容易判断的样本、250例判断困难的样本和1 400例其他情况的样本。

3.1　过滤器的网络部署

在实验中采用基于过滤器串联的网络部署方案，由接收方进行过滤。在内部网络和路由器之间安装一个透明网桥，只对数据链路层数据包中传输层端口号为25的包进行截获，交给过滤器判断，网络内、外的用户感觉不到该过滤器的存在。这种部署方案简单易行，可以充分验证过滤器本身的性能。

3.2　实验结果分析

实验特别关注1 897例垃圾邮件和250例判断困难的正常邮件。将这两类邮件分别以均匀分布和集中式分布的方式安排在样本库中，多次对样本库进行识别后将数据取平均值，结果为：(1) 均匀分布的错误否定(垃圾邮件判断为正常邮件)率为2.83%，错误肯定(正常邮件判断为垃圾邮件)率为3.05%；(2) 集中式分布的错误否定率为4.23%，错误肯定率为4.74%。其中的一次实验结果如图4、图5所示。实验结果表明：在均匀分布的情况下，错误否定率和错误肯定率保持了较低的水平，过滤器性能稳定；在集中式分布的情况下，过滤器两次遭遇人为安排的密集的垃圾邮件样本和判断困难的正常邮件样本，错误否定率和错误肯定率出现波动，但由于抗体网络能有效地通过基于克隆选择、变异机制的无监督竞争学习，识别新出现的抗原，过滤

器的性能很快得到恢复，表现出较强的自适应能力。

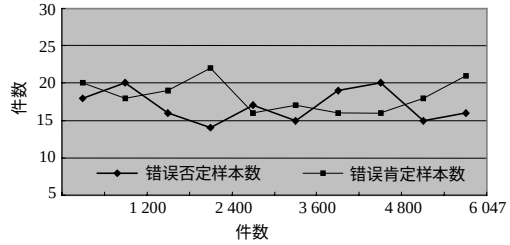

图4 平均分布的测试结果

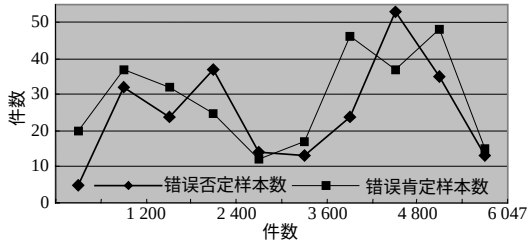

图5 集中式分布的测试结果

4 对抗体网络的几点讨论

本文将文献[6]提出的抗体网络模型用于邮件过滤器的设计。针对这一特定问题，对抗体网络模型作了几点改进，现讨论如下：

(1) 抗体和B细胞的编码。在抗体网络模型中[6]，抗体和B细胞的编码采用无结构的二进制串，不适合用来描述邮件样本。本文将基因库的构造、亲和力的计算、B细胞的变异等操作设计为以特征项为基本单位进行。实验表明，如果采用基于单个字符或基于二进制串的操作，会导致基因库和特征向量中存在大量无意义的特征项，降低识别准确性。

(2) 网络初始结构的形成。文献[6]提出的抗体网络模型，网络需要有一个基于克隆选择的训练、成长过程。本文构造的抗体网络不需要训练过程，初始B细胞种群是随机产生的，不仅提高了效率，而且有利于对整个形态空间的搜索。

(3) B细胞的克隆。在抗体网络模型中，选择细胞进行克隆的依据是抗原浓度ζ_j最大。而ζ_j的大小与过去一段时间内出现的抗原频率有关，因此抗原浓度最大的B细胞不一定对当前抗原有最好的识别效果。对于邮件过滤器这种有较高实时性要求的系统，这种竞争机制会导致网络遭遇新的抗原时，适应时间更长。

(4) B细胞的剪枝。网络中的B细胞如果在一定时间范围内没有识别到抗原，将会从网络中被剪枝并死亡。抗体网络判断B细胞的死亡基于抗原浓度ζ_j和时间t[6]两个参数。本文将其简化为一个参数，即生命值。竞争胜利的B细胞增加生命值，生命值降为0的B细胞死亡，同样实现了“优胜劣汰”。

5 结 束 语

抗体网络作为一种新型的基于免疫原理的神经网络模型，具有结构自动生成，基于克隆选择、变异机制的无监督竞争学习等特征。由于模型提出的时间较晚，国内还缺乏相关研究。相信随着研究的深入，抗体网络会在工程领域得到更广泛的应用。

参 考 文 献

[1] 周伟良, 何 鲲, 曹先彬, 等. 基于一种免疫遗传算法的BP网络设计[J]. 安徽大学学报(自然科学版), 1999, 23(1): 63-66.
[2] 唐 斌, 胡光锐. 基于免疫RBF网络的雷达信号分类识别[J]. 数据采集与处理, 2002, 17(4): 371-375.
[3] 杨淑媛, 焦李成, 刘 芳. 一种免疫径向基多用户检测方法[J]. 西安电子科技大学学报(自然科学版), 2004, 31(2): 209-213.
[4] 付利华, 何华灿. 基于免疫进化规划的一种柔性神经模糊推理系统[J]. 计算机工程与应用, 2002, 38(18): 19-22.
[5] 吕 岗, 谭得健, 赵鹤鸣. 基于免疫算法的前馈神经网络权值设计[J]. 计算机工程与应用, 2002, 38(17): 31-32.
[6] De Castro L N, Von Zuben F J, De Deus J G A. The construction of a boolean competitive neural network using ideas from immunology[J]. Neurocomputing, 2003, 50: 51-85.
[7] 莫宏伟. 人工免疫系统原理与应用[M]. 哈尔滨: 哈尔滨工业大学出版社, 2002.
[8] 李 涛. 计算机免疫学[M]. 北京: 电子工业出版社, 2004.
[9] 曹麒麟, 张千里. 垃圾邮件与反垃圾邮件技术[M]. 北京: 人民邮电出版社, 2003.

编 辑 熊思亮